

Cognitive Schemas and Virtual Reality

Alison McMahan and Warren Buckland

Introduction

In "Orientation in Film Space: A Cognitive Semiotic Approach," Warren Buckland explores "the relation between literal and fictional space in fiction films, as well as the spectator's literal and imaginative orientation in relation to fictional entities (particularly the spectator's location in relation to the fictional geography)." [1] He examines how films simulate spatial and perceptual cues that enable spectators to mentally orient their body in relation to a film's fictional environment. Buckland argues that two levels of "make-believe" are necessary for a fiction to be perceived as fiction: spectators must "take an attitude of make-believe to the events of the fiction... and to the perceptual relations they establish to those events." [2] He refers to the Participatory Thesis, or the Imagined Observer Hypothesis (IOH), which states that, to comprehend the film and to become emotionally engaged, spectators must imagine they are an observer who is "inside the fiction."

As users navigate through the space of a game they must orient themselves. But what space and what type of orientation are we referring to? The two types of spaces to which Buckland refers are real space and its cognitive representation -- imaginative space. Three basic types of orientation exist: absolute, intrinsic, and contextual. Absolute orientation refers to cardinal directions (North, South, East, West), which collectively represent an objective frame of reference within global space, therefore enabling absolute orientation. Intrinsic orientation originates from the properties of objects (e.g., their symmetry), and the frame of reference for intrinsic orientation in a three-dimensional space consists of three axes: up / down, front / back, and left / right. Finally, in contextual orientation, the frame of reference consists of two relative points, for example a speaker and objects or events to which he or she refers. This type of orientation is called deictic. Deictic orientation is therefore egocentric -- that is, organized around the ego (or the body, since the ego is simply the cognitive representation of the body in the mind). Karl Bühler argues that deixis is used not only when the individual is coordinated in real physical space ("ocular deixis") but also in an imagined visual space ("imagination-oriented deixis"). [3]

Kinesthetic Image Schemas

How is it possible for spectators to, first, accept the two-dimensional images on the screen as three-dimensional, and second, to imagine themselves as an observer inside that 3D space? In *The Cognitive Semiotics of Film* [4], Buckland answers this question

by turning to the work of cognitive semanticist George Lakoff, who has posited a series of "Kinesthetic Image Schemas" [5] that structure perceptual input into experiences. These schemata are inherently meaningful because they gain their meaning directly from the body's innate sensory-motor capacities. Kinesthetic Image Schemata therefore represent the body in the mind and are posited as being cognitively real because they are directly motivated or non-arbitrary, and inherently meaningful. Cognitive semantics therefore challenges the dualism (the mind-body problem) of Cartesian philosophy. [6]

Lakoff defines embodiment as "our collective biological capacities and our physical and social experiences as beings functioning in our environment." [7] Kinesthetic image schemata are simple structures that arise from the body -- up-down, back-front, centre-periphery, part-whole, inside-outside, paths, links, forces, and so on. These schemata are directly constrained by the dimensions of the human body. And because the dimensions of the fully-grown human body are shared, fairly uniform and constant, any discussion of conceptual structure in terms of the body does not fall into radical relativism and subjectivism. The structure of our shared bodily experience then becomes the basis for rational, abstract thought by means of image-based schemata and creative strategies such as metaphor and metonymy, which project and extend this structure from the physical domain into the abstract domain of concepts.

As users navigate through the space of a game they must orient themselves. But what space and what type of orientation are we referring to?

Lakoff outlines several kinesthetic image schemata and shows how they determine the structure of abstract conceptual thought. The rest of this paper applies each schema to virtual reality.

The Schemas and Virtual Reality

The Container Schema

The container schema, which structures our fundamental awareness of our bodies, is based on the elements "interior," "boundary," and "exterior." In terms of metaphorical extension, Lakoff notes that the visual field is commonly conceived as a container, since things "come into" and "go out of" sight.

Immersive virtual reality environments, such as CAVEs™, are literal embodiments of our projection of the container schema. CAVEs™ are rooms of 3 x 3 x 3 meters with 3D images projected on floor and walls and using surround sound -- in other words, the "containers" of movie theatre auditorium and movie screen

sense of projection into film space. 3D navigable spaces are much further along the continuum between simulation and copy (though of course we are nowhere near achieving the "ideal" of the holodeck). 3D modeling is designed to further our sense of projection into a volumetric space. In polygon-only systems used with most 3D VR environments, these include a mesh level of detail, scale, height maps, textures, colors, lighting, and depth cueing. [8] Often the worlds projected in 3D are themselves in the shape of containers: how many of us have modeled a dome for a sky over a flat grid for a terrain? Or hidden a little world inside a spherically shaped sky? And so, the 2D flat screen of the cinema in a 6-wall cave or with a VR helmet is turned into a screen-as-container; and within that physical container exist a multiplicity of graphically rendered containers, within which the action of the game or VR environment takes place.

While film spectators imaginatively construct screen space in their mind using the container schema and limited cues from the film, a range of 3D computer software from CAD programs to immersive environments such as CAVEs™ enable spectators to actually "walk" through a space.

have been blended, overlapped, collapsed into each other. While film spectators imaginatively construct screen space in their mind using the container schema and limited cues from the film, a range of 3D computer software from CAD programs to immersive environments such as CAVEs™ enable spectators to actually "walk" through a space. By the same token, someone, such as an architect, who has a mental schema of a room in their mind and wants to make it available to others can use the same programs to create a version of that space in the computer. First person or egocentric perspectives of VR environments, such as CAVEs™, eliminate one level of container projection demanded by the viewer of the film. Instead of having to translate the two-dimensional screen of the film to a three dimensional mental space, as the cinema spectator does, VR users simply steps into a CAVE™, turns on the computer, and puts on their 3D goggles.

In film, the 2D screen simulates a 3D space through various cinematic tools used to give the spectator a

The Part-Whole Schema

The part-whole schema is based on our perception of our bodies as wholes made up of parts. As an example of metaphorical extension, a narrative film is understood to be a whole made up of parts -- shots and scenes. Machinimists, that is, animators who make films using 3D game engines such as *Quake*, *Unreal*, or *Half-life*, are evolving such a part-whole schema for 3D spaces by treating machinima films as an outgrowth of cinema, though the opportunities and limitations of Machinima are leading to certain changes. Machinima tutorials, found on websites like www.machinima.com, often advise practitioners to emulate the way parts and whole relate to each other in the cinema, even though this can be incredibly difficult to do in a game editor. For example a 360° pan, or circling camera, difficult to achieve on a film set, is very easy to do in Machinima and as a result tends to be overused. Slow motion ("slomo") is necessary to make action sequences read clearly, but in films it is

usually used for emotional emphasis (and it is generally avoided in customized game levels). Fades-to-black are used in between scene changes, not to indicate the passage of time as it does in traditional films but to cover up the fact that a new digital map is being loaded. To a filmmaker this feels like a missed opportunity: what is most attractive about the possibilities of machinima is its promise to escape from the limitations of the classical Hollywood cinematic language, just as the promise of VR is to escape from the limitations of the two dimensional screen. [9]

The Link Schema

Relating the parts to the whole is actually described by Lakoff as a "link" schema. Humans are born attached to their mothers through an umbilical cord, and continue to hold onto, or be held by, family members throughout infancy and childhood. This early experience of being linked becomes a mental schema, so that social and interpersonal relationships are understood in terms of being linked, and divorce is seen as "splitting up."

The part-whole schema is closely aligned to the link schema, since a collection of parts can only be perceived as a whole if the parts are linked together coherently. In immersive technologies the linking schema is most applicable to the interface. It is the interface (usually consisting the screens, the goggles, the glove and/or the wand) that enables the user to bridge the gap between the a-virtual reality of "meat-space" to the virtual reality of "cyberspace".

Center-Periphery Schema

The center-periphery schema is similar to the part-whole schema in that it is based on our awareness of our bodies as having centers (the trunk and the internal organs) and peripheries (limbs, hair, etc.). The center gains its importance from the basic fact the preservation of the body is more important to survival than the peripheries.

Levels of focalization and the center-periphery schema are naturally matched. Focalized levels of narration emphasize the character's direct experience of events. The Exocentric perspective of VR is analogous to "external focalization" in film, and egocentric matches "internal focalization (surface)." However, focaliza-

tion in interactive narrative has an additional layer of complexity: first there is the user her- or himself, a real body in real space, but linked to the imaginary space directly through the interface. For example, in the typical CAVE™ experience, one person wears the glasses that dictate the perspective and orientation of the image of the other people standing close by with glasses on that are not attached to the system. The "helmet" person also has a joystick with which to control the travel through the spaces. No one person has exactly the same perspective as the other.

Then there is the avatar. This can be a visible avatar that the user relates to exocentrically (as in all those over-the-shoulder games such as the *Tomb Raider* series where the user is always one step behind their avatar) or, as is most common in CAVEs™ VRs, egocentrically. Common default settings for avatars in a CAVE VR is a "height" of 6 feet and width of the head is two feet. Other VR environments, such as *NICE*, though always egocentric, allow the user to get a glimpse of their avatar; in *NICE* this happens when the user looks at their reflection in a pond. [10]

Many VR researchers are focused on the effect of varying degrees of agency and anthropomorphism on the user's sense of presence. Kristine L. Nowak and Frank Biocca found in an experiment that varied the level of anthropomorphism in the image of the interactants to have surprising results. People responded the same to the agents whether they thought a human or

Then there is the avatar. This can be a visible avatar that the user relates to exocentrically (as in all those over-the-shoulder games such as the *Tomb Raider* series where the user is always one step behind their avatar) or, as is most common in CAVEs™ VRs, egocentrically. Common default settings for avatars in a CAVE VR is a "height" of 6 feet and width of the head is two feet.

a computer controlled the agent, proving that if it acts human, we relate to it as human. Users who test the less-anthropomorphic avatar images (eyes and mouth floating in space) reported a higher sense of co-presence and social presence than those who interacted with avatars that had no image at all and those who

interacted with highly anthropomorphic images (the researchers concluded that this set up higher expectations that led to reduced presence when these expectations were not met). [11]

Source-Path-Goal Schema

The source-path-goal schema is based on our experience of bodily movement in a particular direction from one point to another along a path. Its many metaphorical extensions include its structuring of one's long-term aims and ambitions, which become "sidetracked" or "blocked" by obstacles.

This is the last schema that Lakoff treats in depth. Buckland points out that narrative trajectories can be understood in terms of the source-path-goal schema. This is true for interactive narratives as well, even though interactive narratives have multiform plot structures, because there is usually one beginning (though some games have multiple beginnings, so that the user doesn't start in the same place each time they play), a multiplicity of stops along the way that can be navigated in different ways, and one or more endings.

Conclusion

This paper is meant as a starting point for further research. By focusing first on applying cognitive -- semiotic theories from film to virtual reality, specifically orientation, we hope to have laid the groundwork for a deeper theorizing of how emotion is conveyed and elicited in interactive narratives, and for its application in VREs. [12]

Alison McMahan, Filmmaker, VR Producer, Scholar,
Santa Ana, CA 92705
Warren Buckland, Associate Professor, Chapman
University, One University Drive, Orange, CA 92866

Focalized levels of narration emphasize the character's direct experience of events.

References

- [1] Warren Buckland, "Orientation in Film Space: A Cognitive Semiotic Approach" in *Recherches en communication* 19 (2003), p. 98
- [2] Warren Buckland, "Orientation in Film Space: A Cognitive Semiotic Approach" in *Recherches en communication* 19 (2003), p. 89

[3] K. Bühl, *Theory of Language*, Donald Fraser Goodwin, transl. (John Benjamins: Amsterdam, 1990), Chapter 8

[4] Warren Buckland, *The Cognitive Semiotics of Film*, (Cambridge University Press: Cambridge, U.K., 2000)

[5] George Lakoff, *Women, Fire and Dangerous Things: What Categories Reveal About the Mind* (University of Chicago Press: Chicago, 1987), Chapter 17

[6] Warren Buckland, *The Cognitive Semiotics of Film*, (Cambridge University Press: Cambridge, U.K., 2000), p. 39-40

[7] George Lakoff, *Women, Fire and Dangerous Things: What Categories Reveal About the Mind* (University of Chicago Press: Chicago, 1987), p. 266-267

[8] Dave Morris and Leo Hartas *Game Art: The Graphic Art of Computer Games* (Watson-Guptill: New York, 2003), p. 140-145

[9] Alison McMahan, *The Films of Tim Burton: Animating Live-action in Hollywood* (Continuum: New York, Forthcoming), Chapter 2

[10] William R. Sherman and Alan B. Craig, *Understanding Virtual Reality: Interface, Application and Design* (Morgan Kaufman Publishers: Amsterdam, 2003), p. 466

[11] Kristine L. Nowak and Frank Biocca, "The Effect of Agency and Anthropomorphism on Users' Sense of Telepresence, Copresence, and Social Presence in Virtual Environments" in *Presence*, 12:5 (2003), p. 481-494

[12] Alison McMahan, "Immersion, Engagement, and Presence: a Method for Analyzing 3-D Video Games" in Mark J.P. Wolf and Bernard Perron (eds.), *Video Game Theory* (Routledge: London and New York: 2003), p. 67-86;

Alison McMahan, (2003, October). "Memesis: A Prototype in Biofeedback and Virtual Reality Narration for CAVEs" in Hal Thwaites (ed.), *VSMM2003: Hybrid Reality: Art, Technology and the Human Factor, Proceedings of the Ninth International Conference on Virtual Systems and Multimedia*, Montreal, Canada (Hexagram Institute: Montreal, 2003), p. 694-702;

Alison McMahan, *The Films of Tim Burton: Animating Live-action in Hollywood*. (Continuum: New York, Forthcoming)